

# HR sistema di previsione termine rapporto lavorativo dipendenti e strategie per l'ottimizzazione

Gabriel Scozzarro

10/1/2022

## 1.0 Introduzione

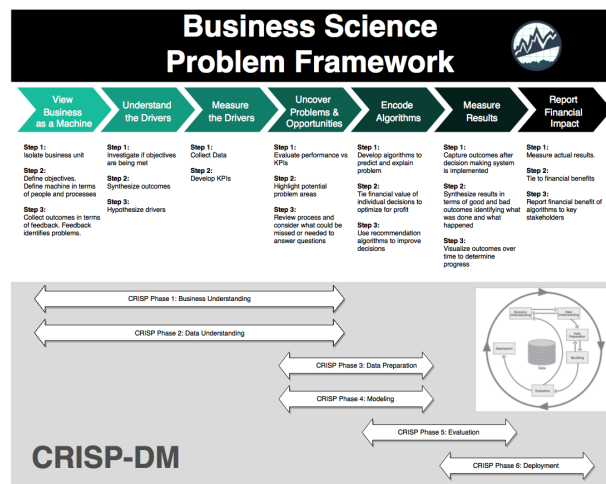
L'interruzione del rapporto lavorativo del personale è definito come il processo attraverso il quale i dipendenti lasciano la forza lavoro – ad esempio per dimissioni per motivi personali o pensionamento e non vengono immediatamente sostituiti. Alcune forme di abbandono sono inevitabili, ad esempio se un dipendente va in pensione o si trasferisce in un'altra città, ma dopo una certa soglia, l'interruzione del rapporto lavorativo può intaccare notevolmente i profitti dell'azienda e la sua cultura. In questo business case è stato realizzato un sistema di predizione dell'interruzione volontaria del rapporto lavorativo basato sui dati forniti dal dipartimento HR dell'azienda. Basandosi sui risultati stimati dal suddetto sistema sono state anche realizzate delle strategie per l'ottimizzazione dell'assegnazione di ore lavorative straordinarie e di incentivi salariali come le stock option.

## 2.0 Data

I dati utilizzati nel presente lavoro sono stati forniti dal dipartimento HR dell'azienda cliente. Contengono informazioni di tutti i dipendenti che hanno lavorato o lavorano nell'azienda.

## 3.0 Strumenti e metodologia

Per questo lavoro è stato utilizzato il linguaggio di analisi e programmazione R attraverso il software RStudio. La metodologia di lavoro utilizzata è la *'Business Science Project Management Framework'* che combina diversi approcci di lavoro, incluso lo standard industriale **CRISP-DM**



## 4.0 Analisi situazione dell'azienda

I dati forniti dall'azienda presentano informazioni riguardanti 1470 dipendenti. E' stata effettuata un'analisi generale per avere una visione di insieme dell'azienda.

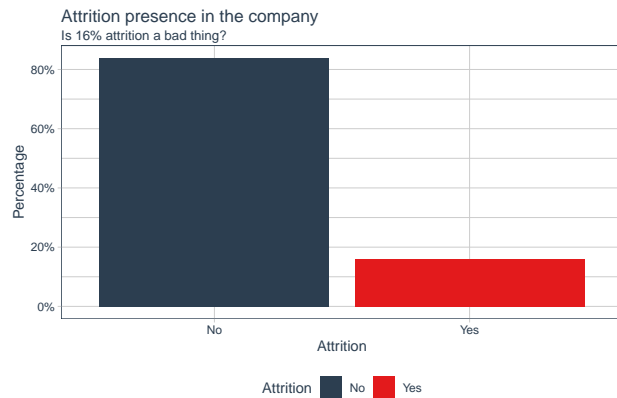


Fig.1

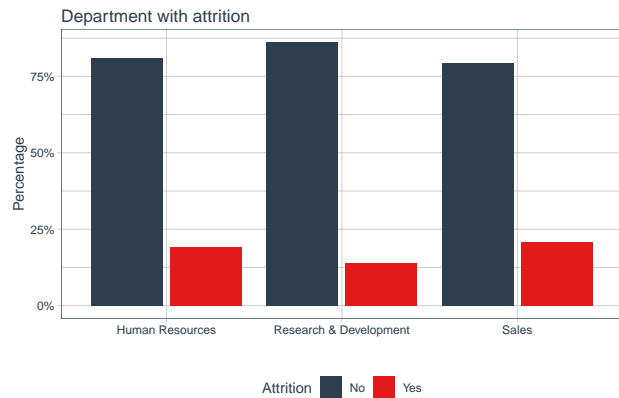


Fig.2

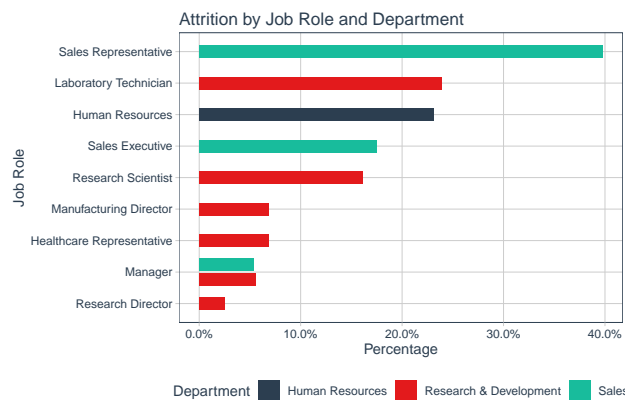


Fig.3

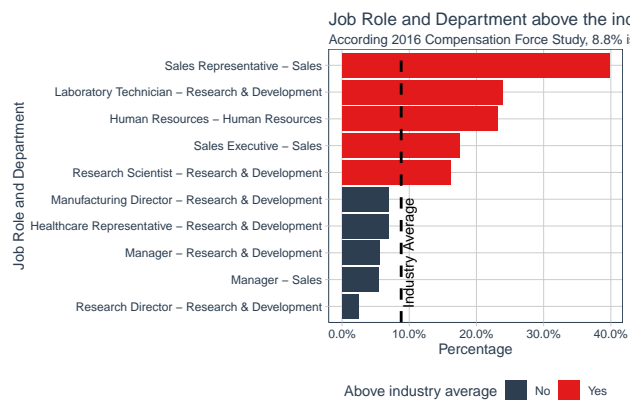
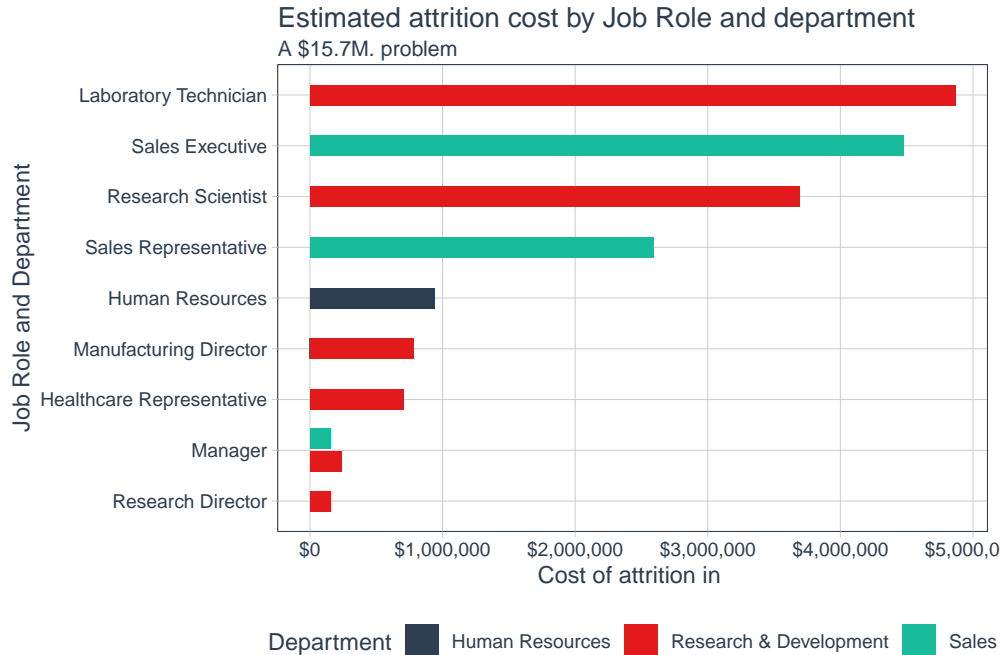


Fig.4

Analizzando in modo generale questi dati abbiamo riscontrato la percentuale di interruzione del rapporto lavorativo in tutta l'azienda. Successivamente, è stato verificato se vi fosse una concentrazione maggiore sia nei diversi dipartimenti sia nelle varie posizioni lavorative. Inoltre, è stata comparata la condizione attuale dell'azienda con la percentuale di interruzioni di rapporti lavorativi media secondo lo standard di riferimento per il settore di appartenenza. Da questa analisi abbiamo riscontrato che:

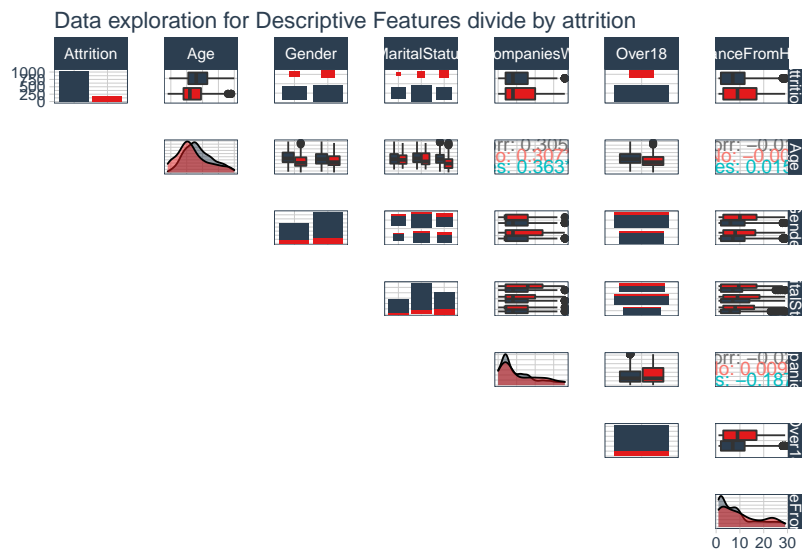
- La percentuale di interruzioni del rapporto lavorativo in tutta l'azienda è il 16% (Fig.1)
- Il dipartimento più afflitto da questa problematica è **Sales** che distacca di pochi punti % il dipartimento di **HR** (Fig.2)
- All'interno dei dipartimenti la posizione lavorativa maggiormente soggetta a interruzioni è **Sales Representative** (Fig.3)
- Confrontando la situazione attuale dell'azienda con la media di riferimento di settore (8.8%) sono 5 le posizioni lavorative su cui intervenire maggiormente. (Fig.4)



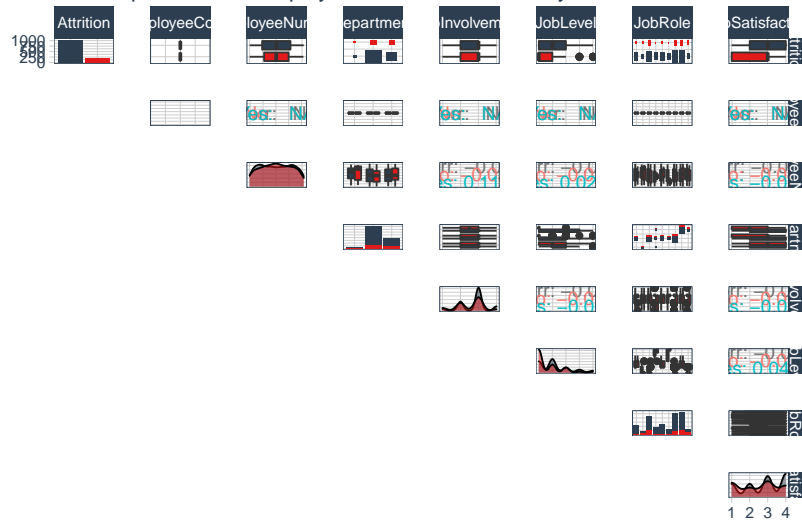
Se traduciamo quanto riscontrato in termini economici riscontriamo un enorme spesa annuale stimata a \$15.7M per l'azienda. Per calcolare questa cifra sono stati ipotizzati i seguenti valori: stipendio medio 80.000, costo di separazione 500, costo assenza dipendente 10.000, costo acquisizione nuovo dipendente 4.900, costo del placement 3.500, giorni lavorativi annuali 240, giorni posizione aperta 40, giorni per mentoring e training 60, efficacia processo di onboarding 0.5, guadagno generato dal dipendente 250.000.

## 5.0 Exploratory Data Analysis (EDA)

Per realizzare il sistema predittivo è stato necessario esplorare le variabili, contenute del database fornito, associate ad ogni dipendente.



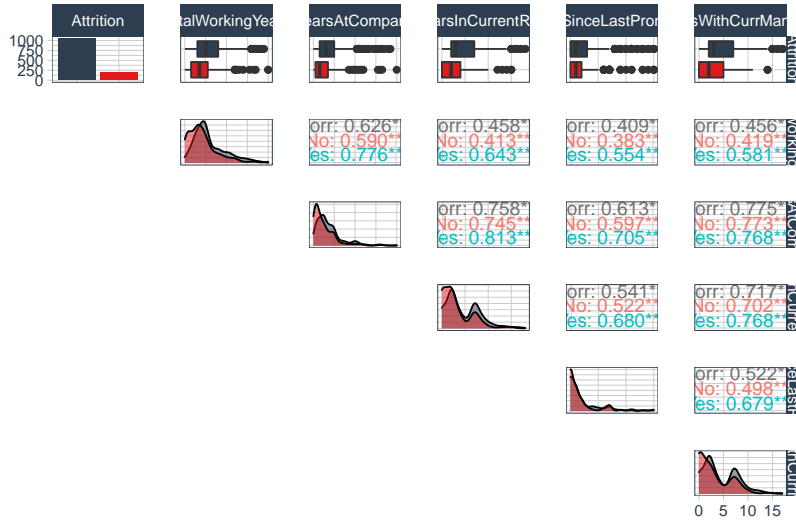
### Data exploration for Employment Features divide by attrition



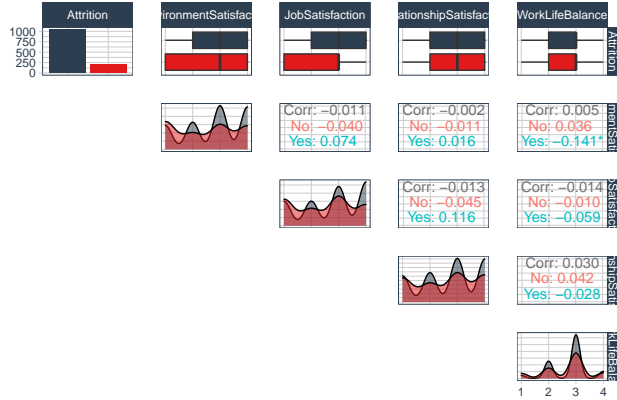
### Data exploration for Compensation Features divide by attrition



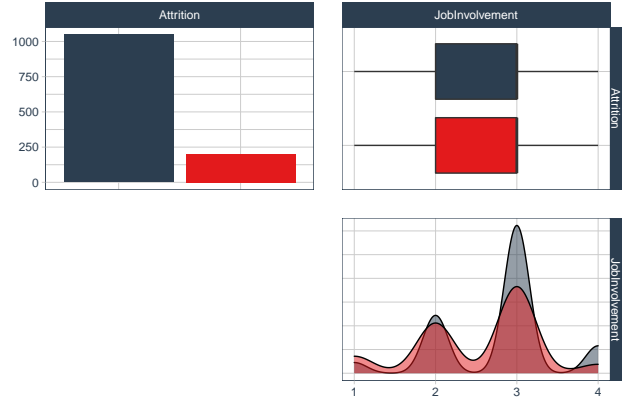
Data exploration for Time based Features divide by attrition



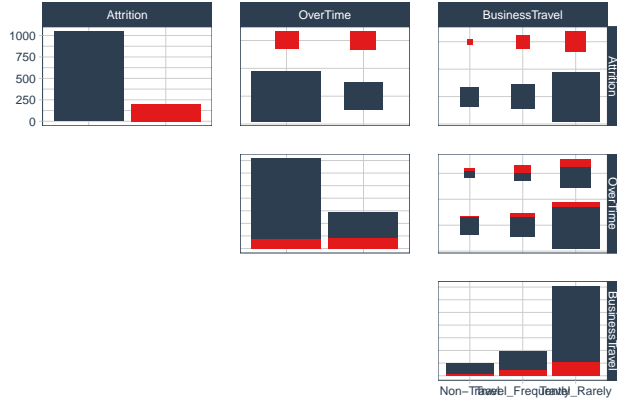
Data exploration for Survey Features divide by attrition



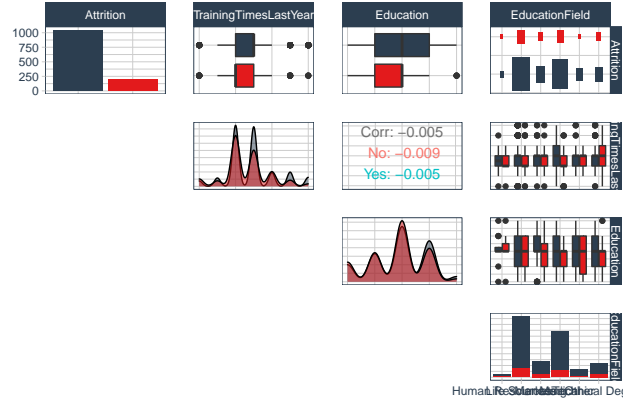
Data exploration for Performance Features divide by attrition



Data exploration for Work-life balance Features divide by attrition



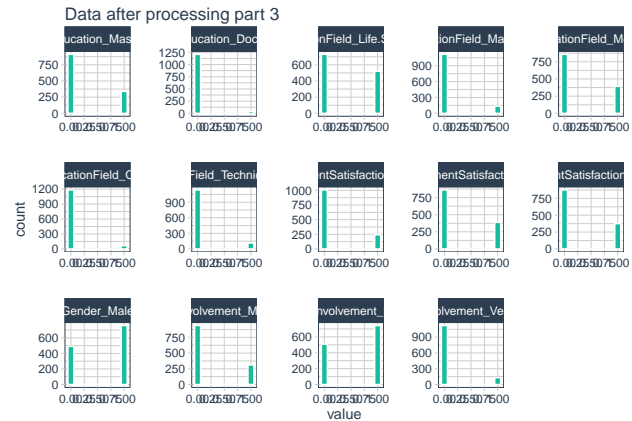
Data exploration for Training and Education Features divide by attrition



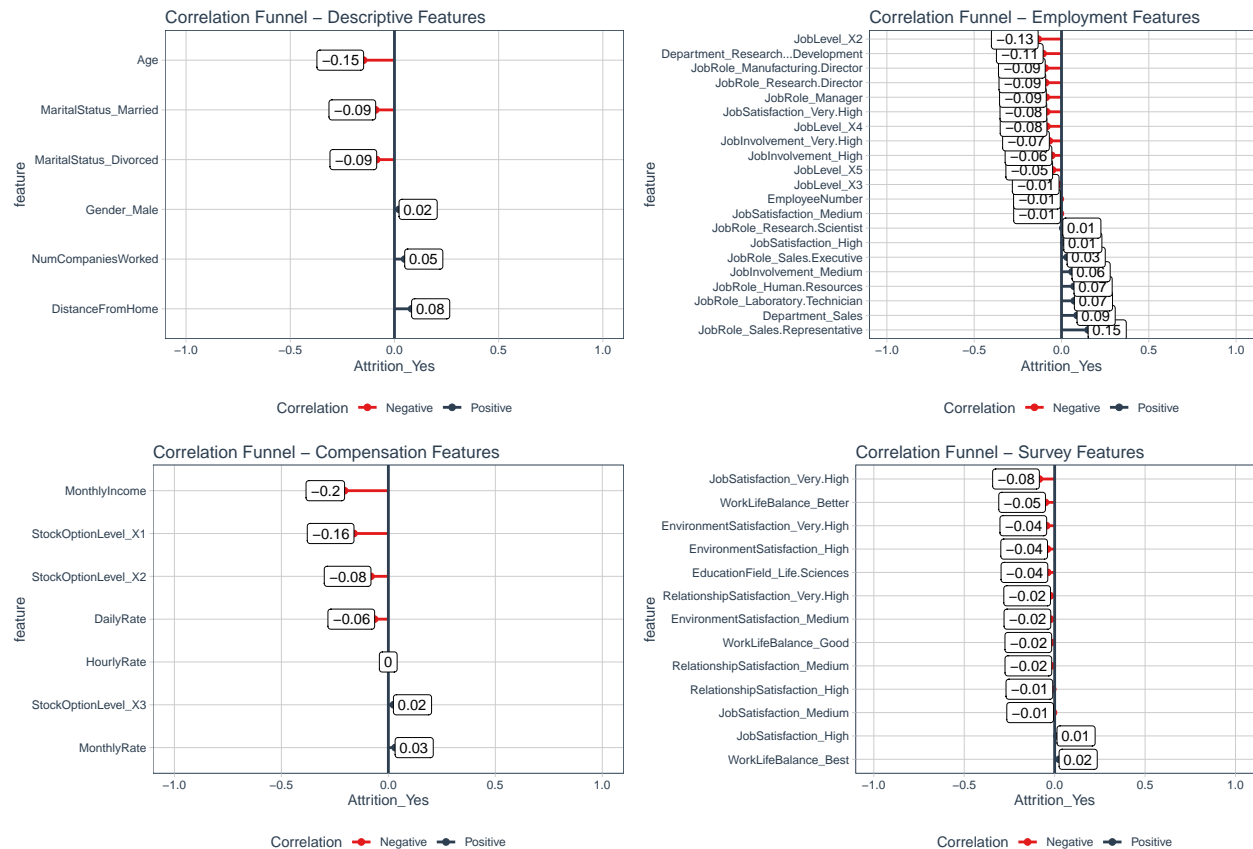
Analizzando le variabili abbiamo riscontrato la necessità di eseguire alcune modifiche in modo da poter essere meglio interpretati. Le modifiche apportate sono:

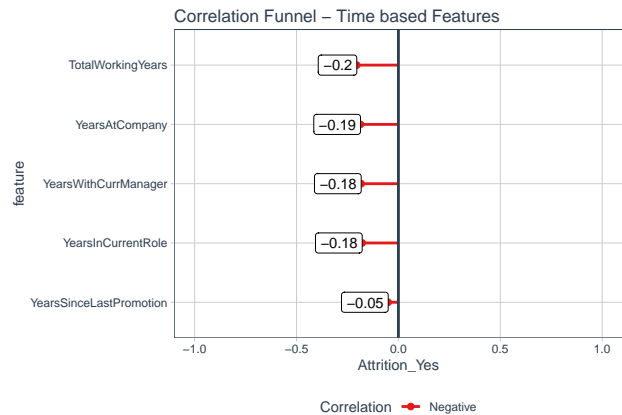
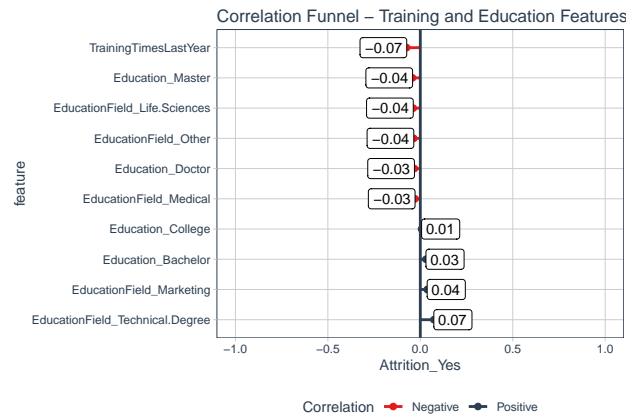
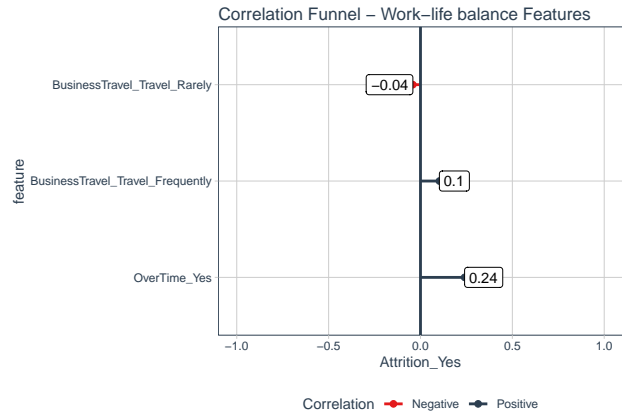
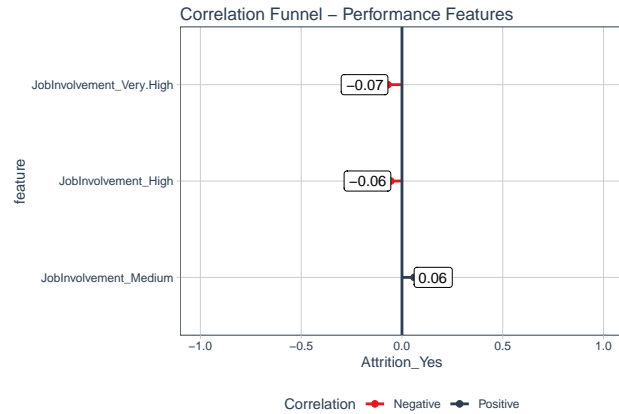
- Rimozione variabile con zero varianza
- Trasformazione di Yeo-Johnson
- Centratatura e normalizzazione variabili numeriche
- Creazione variabili dummy delle variabili nominali





Successivamente a questa lavorazione è stata calcolata la correlazione tra le diverse variabili e l'interuzione del rapporto lavorativo indicata dalla variabile **Attrition**.



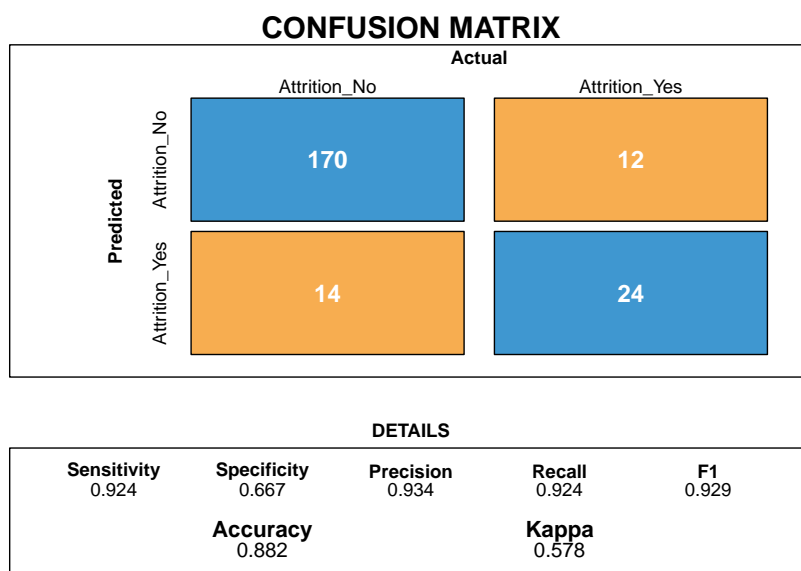
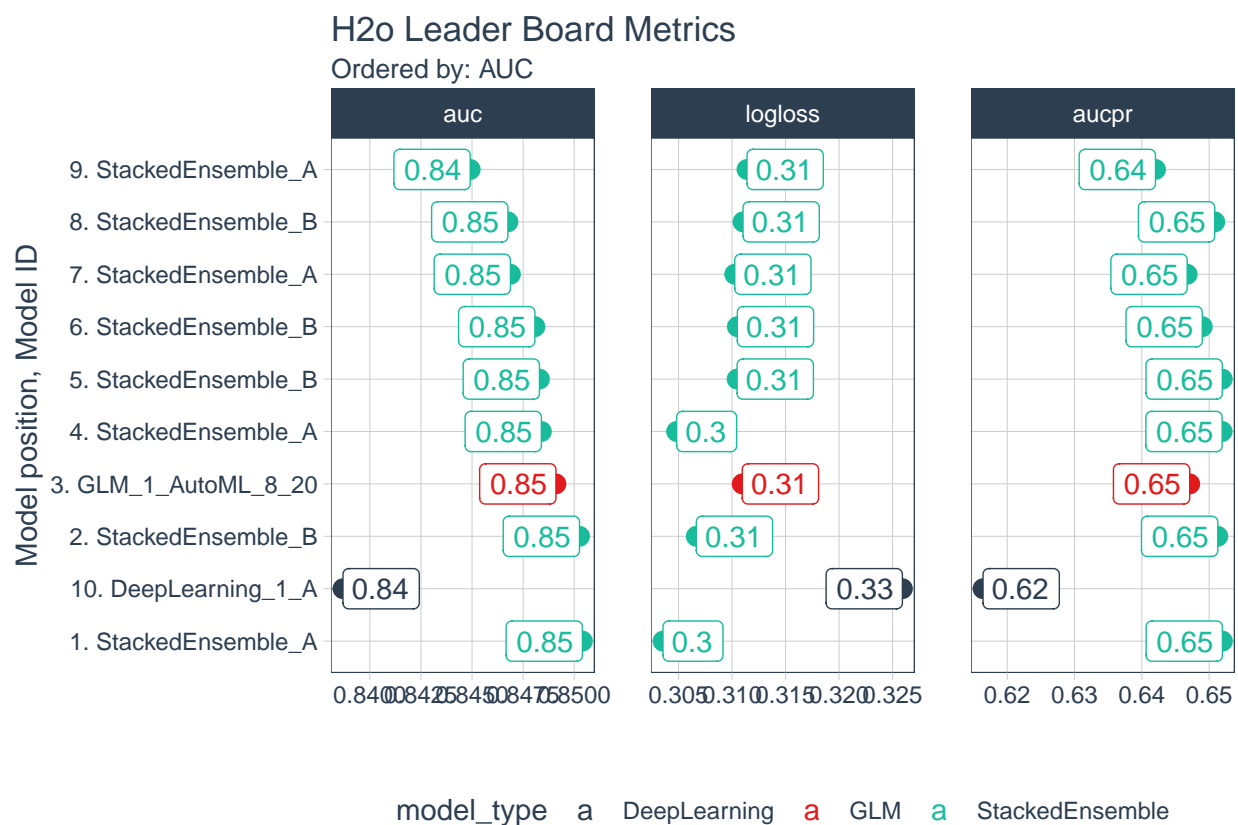


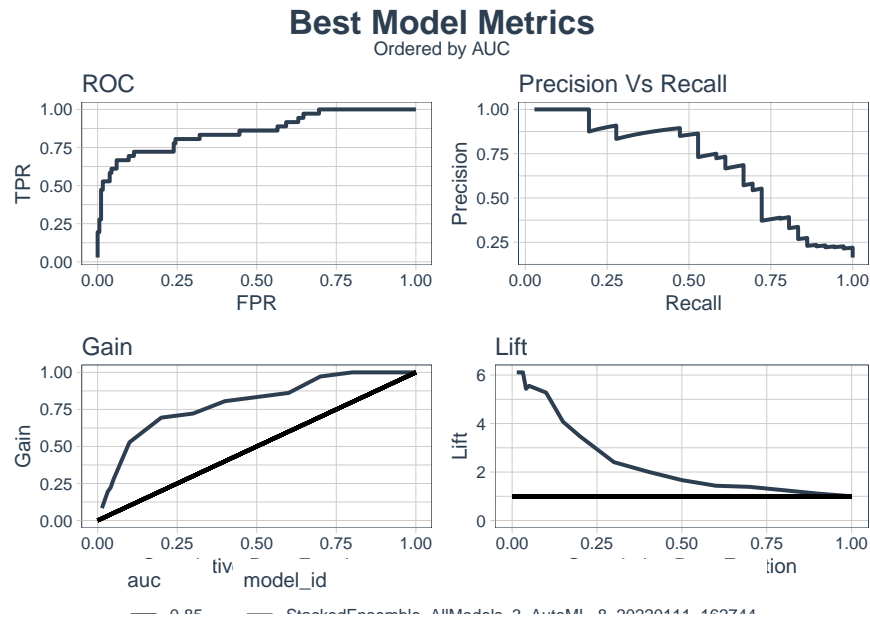
Come evidenziato dai grafici alcuni fattori come la presenza di straordinari, il tipo di incarico lavorativo o la frequenza delle trasferte per motivi di lavoro sono positivamente correlati, in maniera rilevante, con l'interruzione del rapporto lavorativo. In contrapposizione la presenza di incentivi come l'assegnazione di Stock Option, il rating di performance molto alto o l'anzianità di servizio sono negativamente correlati all'interruzione del rapporto lavorativo rappresentando quindi, dei forti driver per il dipendente a rimanere in azienda.



## 6.0 Costruzione del sistema predittivo

Sfruttando il sistema di modellazione automatico del framework H2o.ai è stato sviluppato un modello di *machine learning* in grado di predire se un dipendente interromperà il rapporto lavorativo.

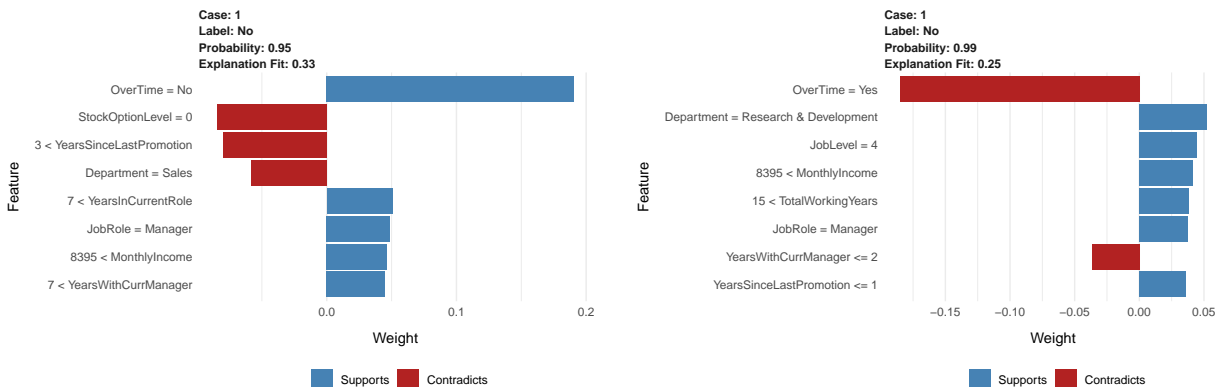


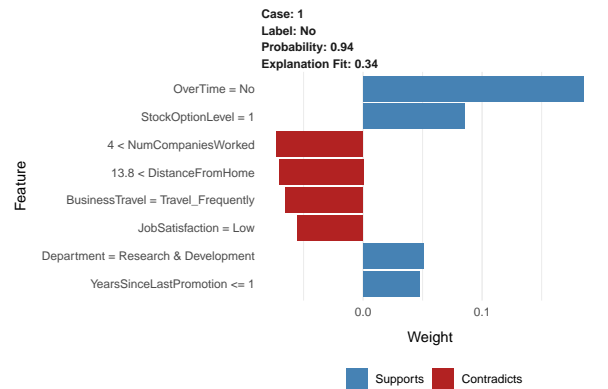
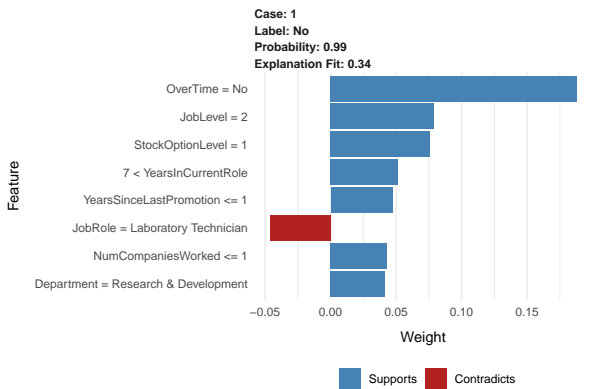
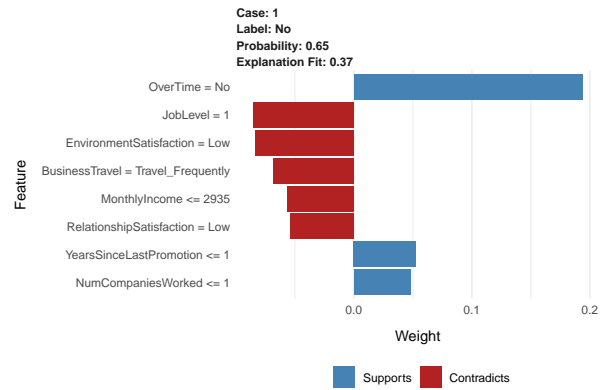
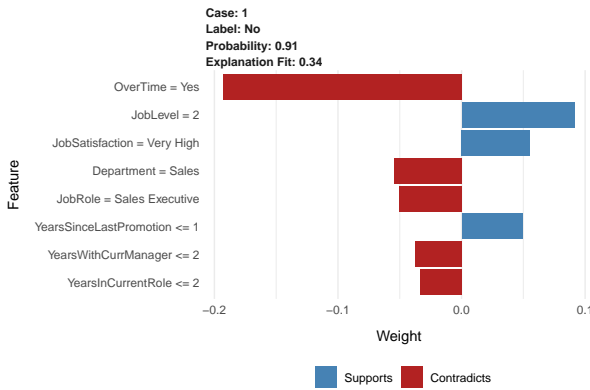


Come mostrato in figura il sistema ha generato diversi modelli che sono stati valutati principalmente utilizzando il punteggio AUC. Il miglior modello ha raggiunto il punteggio AUC di 0.86, un'accuratezza di 0.88, una precisione di 0.93 che indica la suscettibilità del modello ai falsi positivi (dipendenti che vengono indicati come in uscita ma in realtà rimangono) e un richiamo di 0.92 che indica la suscettibilità del modello ai falsi negativi (dipendenti che vengono indicati come stabili ma in realtà terminano il rapporto lavorativo). E' facilmente prevedibile che per questo caso d'uso la capacità minimizzare i falsi negativi abbia un'importanza superiore.

Osservando il grafico del gain è stato possibile riscontrare che senza utilizzare il modello predittivo sviluppato considerando i 220 dipendenti del gruppo di test selezionati casualmente ci aspettiamo di avere 35 dipendenti che termineranno il rapporto lavorativo (media aziendale 16% - pag. 2) se invece venisse utilizzato il sistema predittivo su i primi 10 dipendenti a più altro rischio ne verrebbero già individuati 9, ovvero il 25.7% dei precedenti 35.

Osservando il grafico lift invece, è possibile notare un aumento delle capacità di stima di 5.6 volte (9/1.6) su 10 dipendenti selezionati attraverso il sistema di raccomandazione.

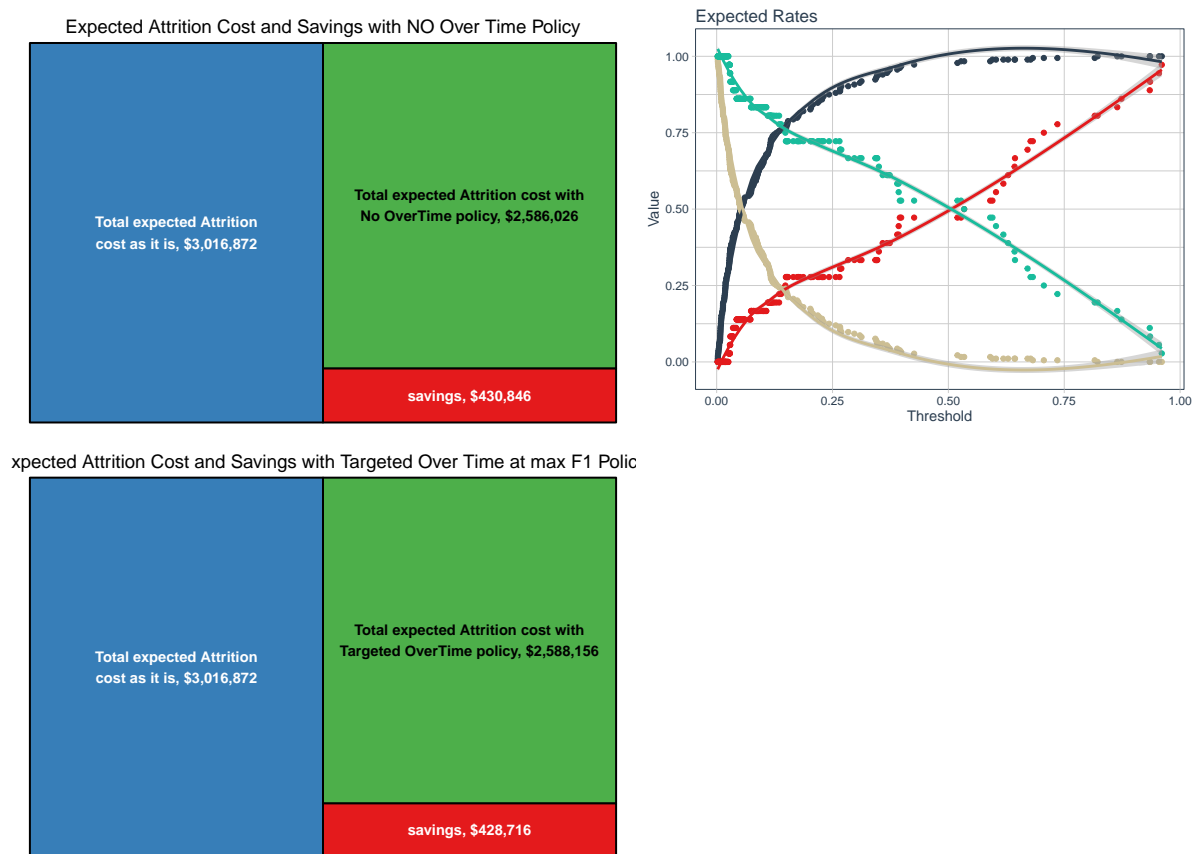




Utilizzando il modello predittivo è stato possibile creare un grafico individuale per ogni dipendente con le 8 variabili che secondo il modello maggiormente supportano o abbattano la probabilità che il dipendente termini il rapporto lavorativo. A titolo di esempio sono stati qui presentati 6 grafici di 6 dipendenti selezionati in modo casuale.

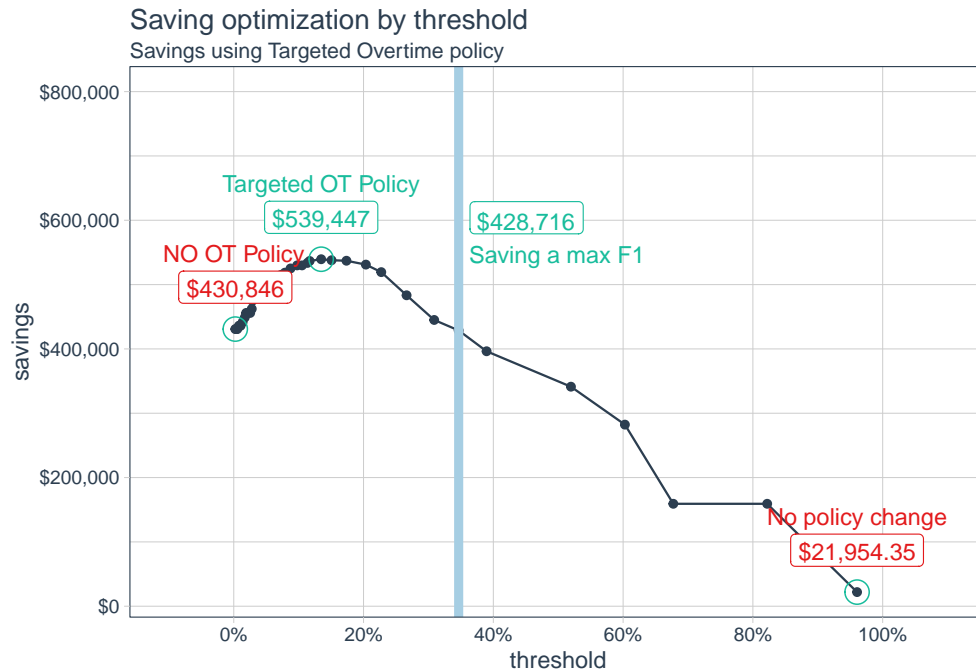
## 7.0 Strategie e outcome stimati

Sfruttando il sistema predittivo sviluppato è stato possibile testare e stimare gli effetti per alcune strategie con l'obiettivo di mitigare gli effetti dei maggiori driver che favoriscono l'interruzione del rapporto lavorativo, come ad esempio l'assegnazione di ore di straordinario o la mancanza di incentivi salariali come le StockOption.

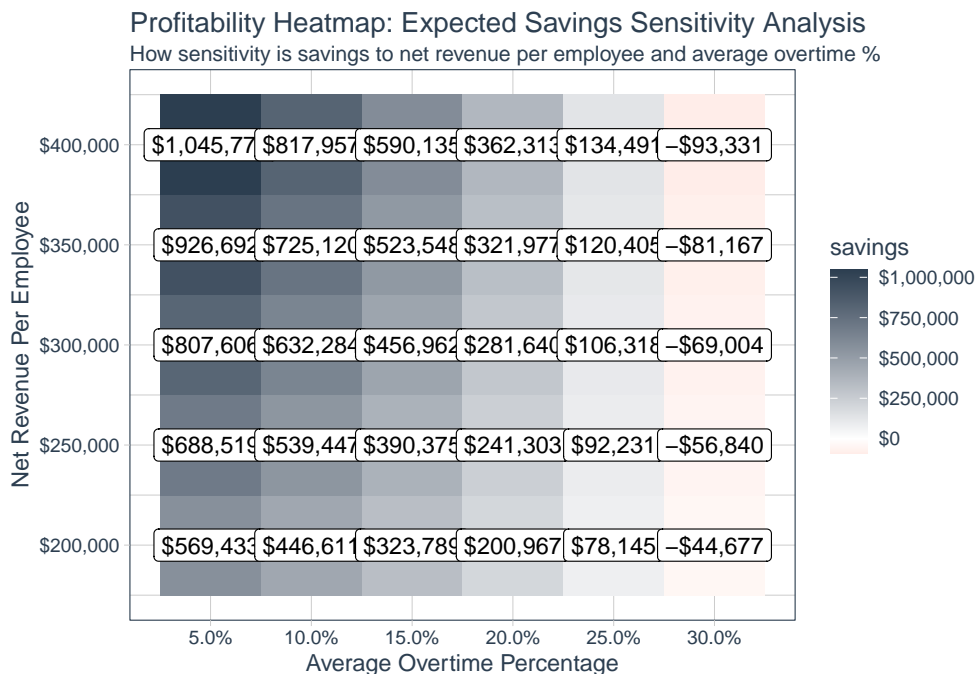


La prima strategia testata grazie al sistema predittivo sviluppato è stata quella di bloccare l'assegnazione di ore di straordinario per tutti i dipendenti in modo da minimizzare il rischio di interruzioni del rapporto lavorativo dovute alle troppe ore. Avremo dunque sia un risparmio sul salario sia un risparmio dei costi legati all'interruzione del rapporto a discapito però della produttività. Come possiamo vedere dal primo grafico il risparmio stimato è di \$430,846 che corrisponde al 14% della spesa "baseline" di \$430,846 calcolata stimando il costo *as it is*.

Successivamente, è stata testata la possibilità di autorizzare le ore di straordinario solo a dipendenti selezionati dal sistema predittivo che non subiscono o in maniera minore, l'influenza negativa delle ore lavorative aggiuntive. In questo caso si cerca di bilanciare i 3 fattori: salario, costo interruzione del rapporto e produttività. Utilizzando i valori di estrapolati dal modello a max F1, che fornisce una misura di quanto bene un classificatore binario può individuare i casi positivi (dato un valore di soglia), è stato simulato di aggiudicare in maniera mirata le ore di straordinario con i risultati mostrati nel terzo grafico. Come possiamo vedere il risparmio stimato è di \$428,716 che corrisponde 14% al rispetto al valore "baseline" calcolato precedentemente.

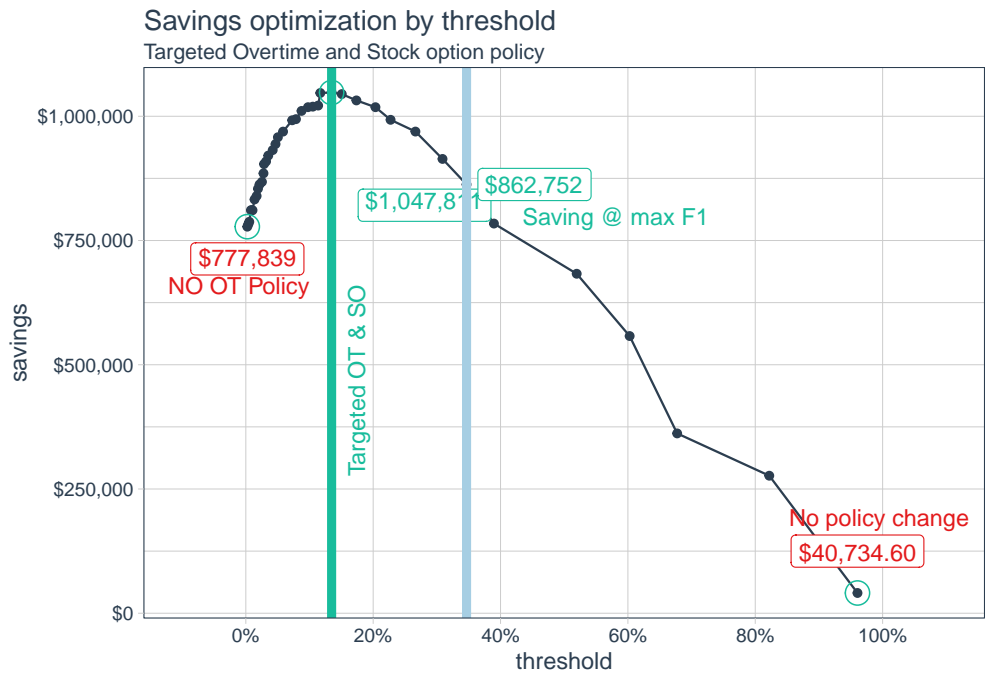


Per massimizzare il risparmio è stato realizzato un sistema di ottimizzazione che andasse a massimizzare il risparmio permettendo al contempo l'assegnazione di ore di straordinario ai dipendenti selezionati dal modello in base allo soglia di classificazione

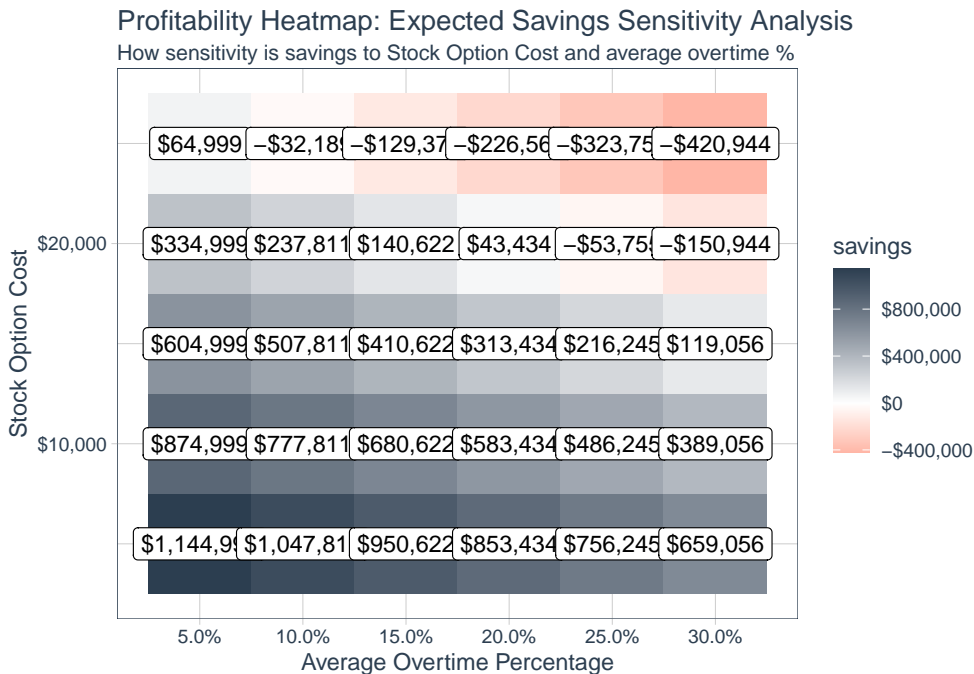


E' stato possibile infine, realizzare un grafico che mostri come la percentuale di straordinari (sul totale di ore contrattualizzate dal dipendente), la produttività e il risparmio stimato rispetto al costo "baseline" delle interruzioni di rapporto lavorativo sia in relazione, quando la strategia di straordinari solo a dipendenti selezionati sia in uso.

In aggiunta alle due strategie precedentemente testate è stato possibile simulare una strategia di incentivo a non interrompere il rapporto lavorativo. Il modello come precedentemente mostrato ha individuato nell'assegnazione di Stock Options un forte incentivo al rimanere.

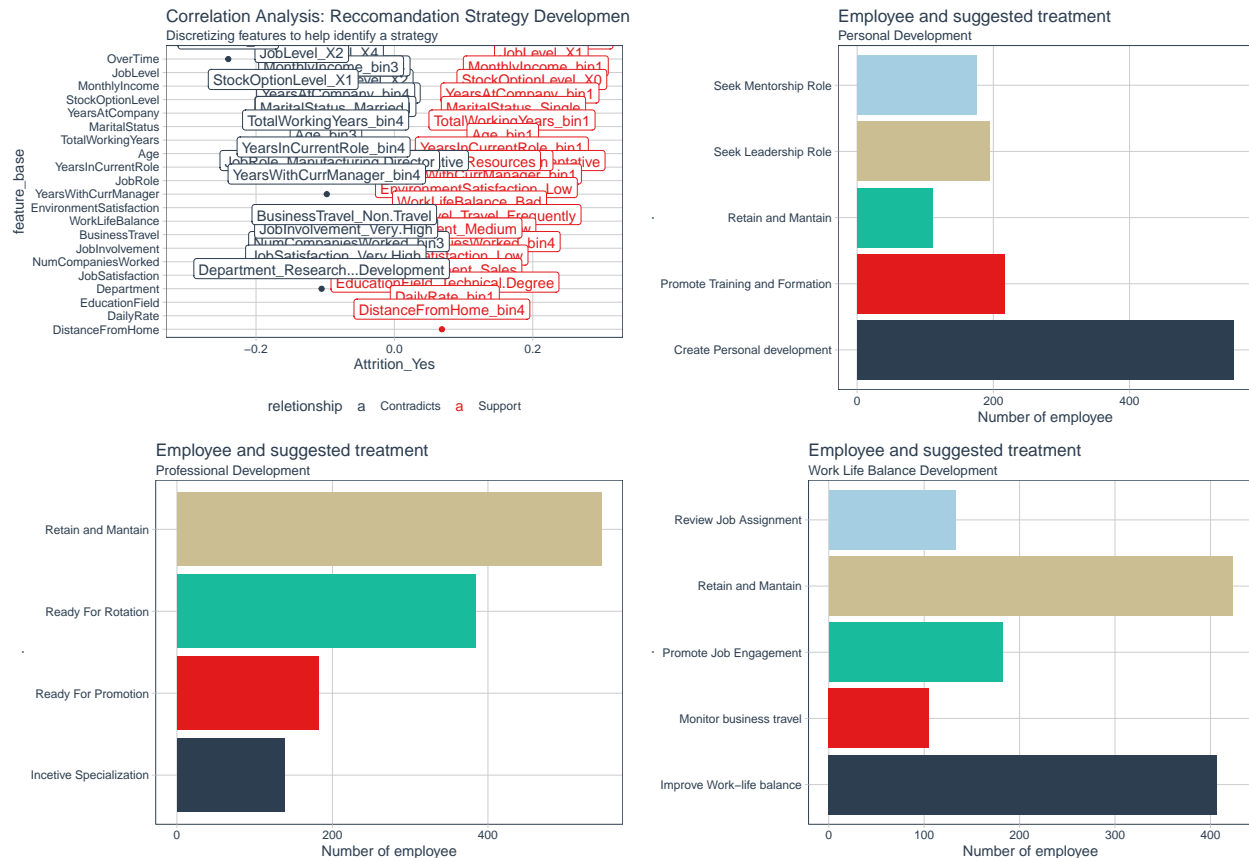


Utilizzando entrambe le strategie di assegnazione mirata sia delle ore di straordinario sia degli incentivi salariali come le Stock Option è stato migliorato notevolmente il valore del risparmio stimato sui costi di interruzione del rapporto lavorativo. Anche in questo caso, è stato possibile creare un grafico che mettesse in relazione sia l'assegnazione delle ore di straordinario, sia quella degli incentivi salariali e il risparmio stimato sui costi di interruzione del rapporto lavorativo.



## 8.0 Strategie personali per dipendente

Come ultima attività è stato sviluppato un sistema di raccomandazione di strategie personali per ogni dipendente, il tutto basato sul medesimo modello matematico.



Negli ultimi 3 grafici è mostrato diviso per categoria (Personale, Professionale, Work-life) gli interventi suggeriti e il numero totale di dipendenti che ne necessitano. E' anche possibile ottenere gli interventi suggeriti anche individualmente utilizzando il numero identificativo del dipendente.